Research Article

A Novel SAR Image Target Recognition Algorithm under Big Data Analysis

Xiang Chen, Xing Wang, You Chen, and Haihan Wang

Aeronautics Engineering College, Air Force Engineering University, Xi’an, China

Correspondence should be addressed to Xiang Chen; 18811506679@163.com

Received 11 August 2021; Accepted 17 September 2021; Published 13 October 2021

Academic Editor: Thippa Reddy G

Copyright © 2021 Xiang Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Synthetic aperture radar (SAR) image target recognition technology is aimed at automatically determining the presence or absence of target information from the input SAR image and improving the efficiency and accuracy of SAR image interpretation. Based on big data analysis, dirty data is removed, clean data is returned, and standardized processing of SAR image data is realized. At the same time, by establishing a statistical model of coherent speckles, the convolutional autoencoder is used to denoise the SAR image. Finally, the network model modified by softmax cross-entropy loss and Fisher loss is used for automatic target recognition. Based on the MSTAR data set, two scene graphs containing the target synthesized by the background image and the target slice are used for experiments. Several comparative experiments have verified the effectiveness of the classification and recognition model in this paper.

1. Introduction

In recent years, SAR image automatic target recognition technology (SARATR) has been widely used [1] and has formed a fixed three-level flow process: detection, identification, and classification [2]. The detection module is mainly based on the detection algorithm to obtain slices containing the SAR image target [3]; the discrimination module eliminates the false alarm value for the target slice [4]; the classification module selects the best decision mechanism to judge the category. With the rapid development of deep learning [5] in machine vision, its models can independently learn the internal laws of data from massive sample data and have strong feature expression ability. For the classification and recognition task of natural images, researchers have proposed some very typical deep learning network models, such as AlexNet [6], VGG [7], Googlenet [8], ResNet [9], and DenseNet [10]. These network models have obtained excellent results in the large-scale visual recognition challenge.

In order to accelerate the realization of efficient classification and recognition of deep learning in SAR targets, many scholars at home and abroad have carried out a series of studies on it. Housseini et al. [11] proposed to learn the convolution kernel and bias based on the convolutional autoencoder (CAE). This method guarantees a high recognition rate; the recognition speed is approximately 27 times that of the CNN architecture alone. Wang et al. [12] studied the influence of coherent speckles in SAR images on CNN for SAR target recognition. On this basis, they proposed a bipolar coupled CNN structure. They firstly used the denoising subnetwork to denoise and then learned the residual speckle characteristics and target information through the classification subnetwork. This structure can improve the noise robustness of the network. Wagner [13] combined convolutional neural network and support vector machine to classify 10 types of targets in the MSTAR database and finally got a recognition rate of 98.6%. Chen et al. [14] took VGG as the reference object and designed a 5-layer CNN network (A-ConvNet) with a large convolution kernel and full convolution mapping, which was cut from the target slice of the MSTAR data set and used as a training sample. The recognition rate reaches 99% in the standard environment and 96% in the 17°-30° extended environment. Zelnio et al. [15] used an 8-layer CNN structure to classify 10 types
of vehicle targets, and its network structure design was similar to AlexNet, with a final recognition rate of 92.3%.

In practical application scenes, the accuracy of SAR image recognition is easily affected by noise interference and multiscale target transformation. In order to obtain effective battlefield situation awareness information, it is of great significance to study advanced SARATR technology. In this paper, a novel model framework based on deep learning is proposed to study the automatic target recognition technology of SAR image. Firstly, preprocess SAR image data through big data analysis methods. At the same time, by establishing a statistical model of coherent speckles, the convolutional autoencoder is used to denoise the SAR image. Finally, the network model modified by softmax cross-entropy loss and Fisher loss is used for automatic target recognition. Based on the MSTAR data set, experimental results show that the proposed method has better recognition performance than other traditional algorithms.

2. SAR Image Target Recognition Method

2.1. SAR Image Data Preprocessing. The big data analysis method is used to preprocess the SAR image data. The standardization process [16, 17] includes three main steps:

(1) Image Data Analysis. Firstly, analyze the SAR image data sample, obtain the attributes and various values of each basic data, and analyze the data quality based on the big data analysis according to the correctness and consistency of the SAR image data.

(2) Define workflow and conversion rules. For dirty data, build standardized data processing steps and conversion rules based on the quality and scale of the data. For data with pattern-level problems, it is necessary to specify a matching and query language to generate its processing code

(3) Reflow clean data. First, back up the SAR image data to be processed, and implement standardized processing on the SAR image data through the final workflow and conversion rules. Store the processed SAR image data in the original data source, and delete the original data entry

2.2. SAR Image Coherent Speckle Noise Statistical Model. Due to the coherent interference of radar waves reflected by many basic scatterers, the SAR image itself is affected by coherent speckle noise, which will reduce the spatial resolution of the image, blur the edge and texture characteristics of the image [18], and greatly affect the interpretation of the SAR image work [19]. Therefore, modeling and suppressing coherent speckle noise in SAR images is an important part of target recognition.

Coherent speckle noise modeling is mainly based on the assumption of fully developed coherent speckle, which can usually be regarded as a random multiplicative noise:

\[ Y = X.I. \]  

Among them, \( Y \) represents the actual observed image intensity (including coherent speckle noise); \( X \) is the noise-free image intensity (the actual situation does not exist); \( I \) represents the random multiplicative coherent speckle noise intensity, which is statistically independent of \( X \).

When the multiview coherent speckle noise satisfies the gamma distribution, the corresponding probability density function [20] is as follows:

\[ P_I(I) = \frac{I^{L-1}e^{-LI}}{\Gamma(L)}. \]  

Among them, \( L \) is the number of sights; \( I \) is the intensity of coherent speckle noise.
3. SAR Target Recognition

This research proposes a SAR image target classification feature learning framework based on CNN and trains the network model through classification tasks. The trained network performs feature extraction on the region of interest in the SAR image and finally inputs the feature information into the classifier to achieve target classification.

3.1. SAR Image Speckle Suppression Based on CAE Network.

Convolutional autoencoder (CAE) [21–23] replaces the fully connected neurons in the autoencoder with convolutional neurons. The principle structure of CAE is shown in Figure 1. CAE consists of two parts: encoder and decoder. An encoder is used to extract image features, and a decoder is used to generate images. In the coding process, feature information of the image is saved and redundant information is
deleted [24]. As the capability of feature extraction is enhanced with the deepening of network layers, when deep CAE network is applied to image tasks, the layers of encoder and decoder are usually deepened at the same time to extract, encode, and decode input features.

After reflooding clean data through the big data analysis method, according to the SAR image coherent speckle noise statistical model [25] in Section 2.2, the coherent speckle noise of the SAR image data needs to be processed [19].

The coherent speckle suppression network used in this article combines the structural characteristics of CAE, as shown in Figure 2. First, the noise suppression network needs to be trained with the coherent speckle statistical model. The noise suppression network includes an encoding layer and a decoding layer. The encoding layer contains four identical convolutional layers and pooling layers. The compression factor is 16 after the entire encoding process. The corresponding decoding layer also includes four deconvolution layers and the upper layer combination structure. The decoding layer is finally enlarged by 16 times, and the original image is restored.

As SAR image coherent spot generally unable to get the noise level, in order to make the network generalize performance, this study adopts mixed noise simulation coherent spot data of network training. The cost function of fine-tuning optimization is as follows:

\[
L_{\text{CoAE}}(x, \hat{x}) = \lambda \sum_{i=1}^{n} |x_i - \hat{x}_i| + (1 - \lambda) \sum_{i=1}^{n} (x_i - \hat{x}_i)^2, \tag{3}
\]

where \(\lambda\) is the super parameter that controls the mixing loss, \(x\) represents the image with noise, \(\hat{x}\) represents the estimated image without noise, and \(L_{\text{CoAE}}\) represents the reconstruction loss of the entire network. The weights are fine-tuned through the softmax classification layer, which not only effectively avoids network overfitting, but also enables the model to better adapt to the SAR image classification and denoising task.

3.2. Target Feature Extraction. In order to perform target recognition, first, need to extract relevant features to distinguish between different targets and then obtain a list of regions of interest (ROI). This paper uses the TP-CFAR algorithm to extract target feature information to provide support for the training of the classifier. The main process of the algorithm is shown in Figure 3.

3.2.1. DP-CFAR Detection Algorithm. Double parameter constant false alarm rate (DP-CFAR) is a kind of constant false alarm probability algorithm. The DP-CFAR algorithm [26, 27] refers to the use of mathematical statistics theory to estimate the parameters of the detection model while keeping the target false alarm rate unchanged, which not only reduces the calculation amount of the algorithm, but also adaptively adjusts the threshold in complex background environment.

Figure 4 shows the main steps of the DP-CFAR algorithm. When the DP-CFAR algorithm is used to detect the target image, sliding window method is usually used to traverse the detection and estimate the target scale [22]. In the actual environment, the target detection background environment of SAR images is more complicated, and it is necessary to design a DP-CFAR detection algorithm that adaptively adjusts the threshold. This paper uses the sliding window method to traverse the pixels to detect the background and the target. The size of the sliding window is estimated according to the estimated target scale acquisition. The detection algorithm first calculates the clutter statistical characteristics of the background area and establishes the background clutter distribution probability density function.

Assume that the statistical distribution of clutter pixel values in the background window is Gaussian. Thus, the false alarm rate of two-parameter CFAR can be calculated as follows:

\[
P_{\text{fa}} = \int_{VT}^{\infty} P(x|H_0)dx = \int_{VT}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{(x-\mu)^2/2\sigma^2} dx, \tag{4}
\]

where \(\sigma\) is the standard deviation of the Gaussian distribution, \(\mu\) is the mean, and \(x_i\) is the value of pixels in the clutter background.

Let \(t = x - \mu/\sigma^2\), then the above equation can be written as follows:

\[
P_{\text{fa}} = \int_{t}^{\infty} p(t|\mu_0)dx. \tag{5}
\]

When the false alarm rate of dual-parameter CFAR is obtained, according to the formula \(\Phi(\alpha) = 1 - P_{\text{fa}}\), the threshold factor \(\alpha\) can be calculated. The threshold corresponding to two-parameter CFAR is \(V_T\).

\[
V_T = \mu + \alpha \sigma. \tag{6}
\]
Formula (6) also satisfies the condition of constant false alarm rate. The estimated values of Gaussian distribution statistics $\mu$ and $\sigma$ are as follows:

\[
\hat{\mu} = \frac{1}{N} \sum_{i=1}^{N} x_i,
\]

\[
\hat{\sigma} = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \hat{\mu})^2}.
\]

In the formula, $\hat{\mu}$ represents the mean estimation; $\hat{\sigma}$ represents the standard deviation estimate of the Gaussian distribution. The final discriminant formula of the DP-CFAR detector is as follows:

\[
V_T = \frac{x_i - \hat{\mu}}{\hat{\sigma}} \begin{cases} \geq \alpha, & x_i \text{ is target,} \\ < \alpha, & x_i \text{ is background.} \end{cases}
\]

The sliding window method is used to adaptively estimate the threshold of target detection, and finally, the classification of the pixels to be detected is judged. The judgment standards are as follows:

- When $P(x|w_0) > P(t|w_0)$, $x$ is background; when $P(x|w_0) \leq P(t|w_0)$, $x$ is target.
After detecting the image, the DP-CFAR algorithm outputs a logical matrix with the same size as the image to be detected. The pixel value of the image mask is 0 for background and 1 for target. In order to remove the isolated false alarm points, the calculated logic matrix and mask image are used for processing. In this paper, the image mask is first corroded to cut off the small connections between adjacent targets in the mask; then, expansion is performed to make up for the small holes of the target in the image mask; finally, the prior knowledge is used to estimate the size of the connected components where the target may exist, and the small connected components are removed.

### 3.2.2. Generate ROI Based on NPA Algorithm

1. **Neighbor Aggregation**. In order to avoid the phenomenon of “increased batches,” the influence of noise may cause pixels with pixel values below the DP-CFAR detection threshold to exist in the target area, resulting in a target being detected as several small targets. In this paper, the nearest neighbor point pixel aggregation (NPA) algorithm is used for target aggregation. First, define the set of neighborhood points:

   \[ N_i = \left\{ s | D(i, j) \leq \sqrt{M^2 + P^2} \right\}, \tag{9} \]

   where \( S \) is the set of spatial neighborhoods of pixel points \((x_i, y_i)\), \( D(i, j) \) is the Euclidean distance between pixel points \((x_i, y_i)\) and \((x_j, y_j)\). According to the prior information of the target, \( M \) is the number of pixels occupied by the length of the target; \( P \) is the number of pixels occupied by the target width, and the number of pixels occupied by the largest scale of the target is usually taken.

2. This paper uses NPA algorithm to achieve small target aggregation. The removal of false alarm targets mainly uses the target area generated by the aggregation and estimates the target size based on prior knowledge. When the pixel exceeds the set threshold, the false target is removed. Finally, according to the position of the center point of the clustering target area, take...
the center point as the center point of the ROI, cut out a fixed size region from the image to be detected, and output the coordinate information of the ROI.

3.3. SAR Image Automatic Target Classification and Recognition. Fisher discriminant analysis (FDA) is a commonly used subspace analysis method. The main idea of the algorithm is that the constraint of calculating the projection axis is to minimize the intraclass divergence and maximize the intersample divergence. Generally speaking, the CNN network is optimized by softmax cross-entropy loss. When the training samples are limited or the diversity is insufficient, the network is easy to overfit, resulting in poor model generalization ability. The intraclass divergence matrix and the interclass divergence matrix of FDA are integrated into the target equation of network optimization, which is used to fine-tune the weight of each network to improve the compactness and class compactness of the feature space in the case of limited training samples, as shown in Figure 5. The separability improves the generalization performance of the entire network. If the loss function of the classification network training only has the softmax classification loss,

\[ L_s = -\frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{k} O(h_k^i = c) \left( w_j^T x^i - \sum_{j=1}^{k} w_j^T \mu^i \right), \]

Define intraclass divergence matrix \( F_w \) and interclass divergence matrix \( F_b \):

\[ F_w = \sum_{k=1}^{C} \sum_{i=1}^{N_c} (\mu_i - \bar{\mu})(\mu_i - \bar{\mu})^T, \]

\[ F_b = \sum_{k=1}^{C} N_c (\bar{\mu} - \bar{\mu}_c)(\bar{\mu} - \bar{\mu}_c)^T, \]

where \( C \) represents the number of CNN training categories, \( N_c \) is the total number of training samples belonging to the \( c \)-th category, \( \mu_i \) represents the \( i \)-th training sample, and \( \bar{\mu} \) is the mean vector of all training samples: \( \bar{\mu} = \sum_{i=1}^{N} u_i / N \). \( \bar{\mu}_c \) is the mean vector of training samples belonging to the \( c \)-type. Then, the FDA algorithm finds the best projection matrix \( F_{FDA} \):

\[ F_{FDA} = \max_w \frac{\text{Tr}(W^T F_b W)}{\text{Tr}(W^T F_w W)}, \]

According to the FDA algorithm, the Fisher loss constraining the interclass and intraclass distance \( L_F \) can be obtained:

\[ L_F = \frac{1}{cN} \sum_{i=1}^{C} \sum_{j=1}^{N_c} O \left( h_k^i = c, h_k^i - \bar{h}_k^2 + \beta h_k^i - \bar{h}_k^2 \right), \]

\[ O \left( h_k^i = c \right) = \begin{cases} 1, & h_k^i = c \text{ is true}, \\ 0, & h_k^i = c \text{ is false}. \end{cases} \]

\[ L = L_s + \lambda L_F. \]
In the formula, $h_i^K$ represents the output of the corresponding feature layer (the first layer before softmax) of the sample $x_i$; $\bar{h}_i^K$ represents the average output of class $C$ samples at the corresponding feature layer; $\bar{h}^K$ represents the average output of all samples in the corresponding feature layer.

Figure 6(a), using the same data set, shows the network model training structure corrected by softmax cross-entropy loss, and Figure 6(b) shows the network model training structure corrected by softmax cross-entropy loss and Fisher loss. The experimental results show that the clustering effect of the same kind of samples in Figure 6(b) is better, and the center spacing of different types of samples is larger, while the sample clustering effect in Figure 6(a) is poor.

4. Experimental Results and Analysis

In order to evaluate the effectiveness of the proposed SAR image target recognition framework, the experimental part of this paper mainly uses the proposed recognition system to recognize the synthetic SAR image of the target scene to be detected and analyzes the experimental results.

4.1. Experimental Data. The experimental data in this section is the MSTAR data set, as shown in Table 1. Due to the high cost of acquiring scene images with a large number of targets, only SAR image slices of 10 types of targets and some background images of the same area are provided in the MSTAR data set. The background image in the MSTAR data set has an imaging elevation angle of 15°, with a total of 100 scene images.

Firstly, preprocess SAR image data through big data analysis methods. Because the background image and the background in the target slice can be regarded as homogeneous regions and both are informed by the same radar, it is recommended to synthesize the scene image containing the target using the target slice and the background in the MSTAR data set.

In this experiment, two scenes to be detected were synthesized, and the synthesis process is shown in Figure 7.

4.2. Experimental Results and Analysis

4.2.1. Speckle Inhibition. According to the scene synthesis method in Figure 7, two different scene graphs were synthesized in this experiment, as shown in Figures 8(a) and 8(b), and the image size of the two scene graphs is $1474 \times 1784$.

Figure 9 shows the denoising results of synthetic scenes 1 and 2 at different noise levels. It can be seen from Figure 9(a) that with the increase of $L$, the noise level also increases. When $L = 1$, the image is in single-view mode, the image is most destroyed, and the fine texture part of the image (the road in the middle of scene 1) is completely covered by the coherent spots. Judging from the speckle suppression images in Figure 9(b), the restored images of the model under different noise levels are roughly the same, and the speckle suppression effect in homogeneous areas is more obvious (the woods below scene 2).
4.2.2. Target Feature Learning and Detection Results. Figure 10 shows the results of target detection in scene 1 and scene 2. The experimental results show that a total of 42 target areas are detected in scene 1, including 7 false alarm targets, and there is no missed detection. In scene 2, a total of 35 targets were detected, including 12 false alarm targets, no missed detection, almost all false alarm targets appeared in the bushes, because the gray value of the bushes is relatively large, and the proportion of bushes in the whole scene is relatively high, and it causes more false alarms.

Figure 11 shows the corresponding image masks output by the two scenes of Figure 10 in the target detection module. It can be seen that the shape of the region of interest at this time is square, and there is overlap between the targets, but in general, different categories are distinguished.

4.2.3. Result of Target Recognition. According to the target feature learning and detection results, a threshold is set for the classifier to identify false alarms. The threshold is 0.5, that is, when the probability of the current category is greater than 0.5, it is judged as a true category; otherwise, it is a false alarm. The target recognition result using the algorithm in this paper is shown in Figure 12(c). The experimental results show that the number of false alarms in scene 1 is 0, and all categories are successfully recognized. The number of false alarms in scene 2 is 1, and only the bush in the lower left corner is identified as target 6, and the recognition effect is good.
In order to verify the effectiveness of the algorithm in this paper, this study analyzes the recognition effect of the CA-CFAR algorithm and the traditional CFAR algorithm, and the experimental results are shown in Figures 12(a) and 12(b). In scene 1, CA-CFAR and the algorithm in this paper performed very well, no false positives and false negatives occurred, and the recognition effect was good. However, the traditional CFAR algorithm has false alarms due to noise interference. In scene 2, the performance of the three algorithms is not perfect. However, the false positive rate of the algorithm proposed in this paper is lower than the other two algorithms, the number of false positives is 2 times, and there is no false negative number, and the effect is better.

5. Summary and Prospect

This paper has carried out in-depth research on the problem of automatic target recognition in SAR images and designed a SAR image classification feature extraction network framework based on deep convolutional self-encoding CNN, which can use the CFAR algorithm to detect a series of regions of interest at the target location, and at the same time, effective classification features can be automatically learned from training data through classification tasks. To solve the problem of poor network overfitting and generalization performance caused by insufficient data diversity, a loss function was designed based on the Fisher criterion, and the weight of the network was fine-adjusted by the loss function, so that the distance between different categories of samples in the feature space of the network mapping was more discrete and the same samples were more clustered. The trained network model can be used as an effective feature extraction for classification.

With the continuous development of deep learning technology, there are many mature end-to-end optimization recognition systems in the field of ordinary image automatic target recognition. How to combine these algorithms with SAR image automatic target recognition is a promising research direction.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References


